Engineers to transform genomic medicine with deep learning startup

July 22 2015

Professor Brendan Frey (centre-right) and colleagues at the University of Toronto Faculty of Applied Science & Engineering. Credit: Roberta Baker/ U of T Engineering

Evolution has altered the human genome over hundreds of thousands of

years—and now humans can do it in a matter of months.

Faster than anyone expected, scientists have discovered how to read and write the DNA code in a living body, using hand-held genome sequencers and gene-editing systems. But knowing how to write is different from knowing what to write. To diagnose and treat [genetic diseases](link), scientists must predict the biological consequences of both existing mutations and those they plan to introduce.

Deep Genomics, a start-up company spun out of research at the University of Toronto is on a mission to predict the consequences of genomic changes by developing new deep learning technologies.

"Our vision is to change the course of genomic medicine," says Brendan Frey, the company's president and CEO, who is also a professor in The Edward S. Rogers Sr. Department of Electrical & Computer Engineering at the University of Toronto and a Senior Fellow of the Canadian Institute for Advanced Research. "We're inventing a new generation of deep learning technologies that can tell us what will happen within a cell when DNA is altered by natural mutations, therapies or even by deliberate gene editing."

Deep Genomics is the only company to combine more than a decade of world-leading expertise in both deep learning and [genome biology](link). "Companies like Google, Facebook and DeepMind have used deep learning to hugely improve image search, speech recognition and text processing. We're doing something very different. The mission of Deep Genomics is to save lives and improve health," says Frey.

Deep Genomics is now releasing its first product, called SPIDEX, which provides information about how hundreds of millions of DNA mutations may alter splicing in the cell, a process that is crucial for normal development. Because errant splicing is behind many diseases and

disorders, including cancers and autism spectrum disorder, SPIDEX has immediate and practical importance for genetic testing and pharmaceutical development. The science validating the SPIDEX tool was described in the January 9, 2015 issue of the journal *Science*.

"The genome contains a catalogue of genetic variation that is our DNA blueprint for health and disease," says Stephen Scherer, director of The Centre for Applied Genomics at SickKids and the McLaughlin Centre at the University of Toronto, a CIFAR Senior Fellow, and an advisor to Deep Genomics. "Brendan has put together a fantastic team of experts in artificial intelligence and genome biology—if anybody can decode this blueprint and harness it to take us into a new era of genomic medicine, they can."

Until now, geneticists have spent decades experimentally identifying and examining mutations within specific genes that can be clearly connected to disease, such as the BRCA-1 and BRCA-2 genes for breast cancer. However, the number of mutations that could lead to disease is vast and most have not been observed before, let alone studied.

These mystery mutations pose an enormous challenge for current genomic diagnosis. Labs send the mutations they've collected to Deep Genomics, and the company uses their proprietary deep learning system, which includes SPIDEX, to 'read' the genome and assess how likely the mutation is to cause a problem. It can also connect the dots between a variant of unknown significance and a variant that has been linked to disease. "Faced with a new mutation that's never been seen before, our system can determine whether it impacts cellular biochemistry in the same way as some other highly dangerous mutation," says Frey.

Deep Genomics is committed to supporting publicly funded efforts to improve human health. "Soon after our Science paper was published, medical researchers, diagnosticians and genome biologists asked us to

create a database to support academic research," says Frey. "The first thing we're doing with the company is releasing this database—that's very important to us."

"Soon, you'll be able to have your genome sequenced cheaply and easily with a device that plugs into your laptop. The technology already exists," explains Frey. "When genomic data is easily accessible to everyone, the big questions are going to be about interpreting the data and providing people with smart options. That's where we come in."

Deep Genomics envisions a future where computers are trusted to predict the outcome of experiments and treatments, long before anyone picks up a test tube. To realize that vision, the company plans to grow its team of data scientists and computational biologists. Deep Genomics will continue to invent new deep learning technologies and work with diagnosticians and biologists to understand the many complex ways that cells interpret DNA, from transcription and splicing to polyadenylation and translation. Building a thorough understanding of these processes has massive implications for genetic testing, pharmaceutical research and development, personalized medicine and improving human longevity.

Provided by University of Toronto